

Text Mining – Sentiment Analysis

JMP PRO The Text Explorer platform is used to explore frequently used words and phrases in unstructured text data such as text found in product reviews, social media posts, comment fields in surveys, incident reports, etc. This guide shows how to perform a sentiment analysis – a methodology that assigns numerical scores to words and phrases with the intent of quantitatively measuring the positive and negative sentiment expressed in the text.

The text data must first be prepared for these analyses. See the **Text Explorer – Describing Unstructured Text Data** guide and the **JMP Help** for information on using the Text Explorer platform to do this.

Summarizing Text Data

Example: Chips.jmp (Help > Sample Data Folder)

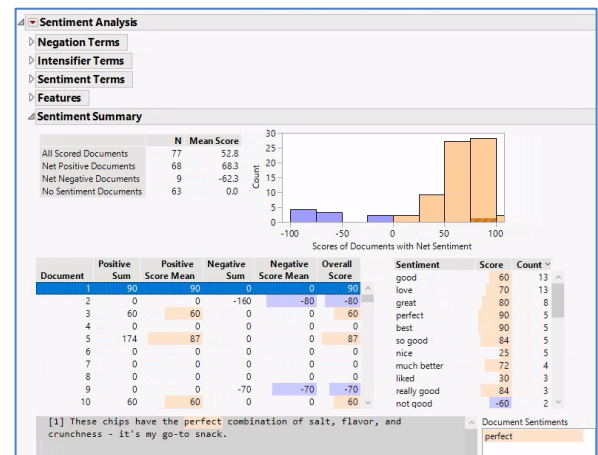
1. From an open JMP® data table, select **Analyze > Text Explorer**.
2. Click on a categorical or unstructured text variable from **Select Columns**, and click **Text Columns** (categorical variables have red or green bars and unstructured text variables have grey lines (≡)). Click **OK**.

Sentiment Analysis

1. Select **Sentiment Analysis** from the top red triangle **Analyze > Text Explorer**.

The following will be displayed in the report:

- A table of the average scores for the positive and negative terms as well as the overall average score for all documents that were scored.
- A histogram visually summarizing all the scores.
- An interactive table showing the scores for each of the documents. Selecting a document will show the text for that document and the words that were used for scoring.
- An interactive table where the words used for scoring can be selected and all the documents that used those words will be displayed.



- The default dictionary used for the analysis will be summarized in a set of tables (**Sentiment Terms**, **Intensifier Terms**, **Negation Terms**) where the scores for each word can be viewed and modified.
- JMP will display a set of **Possible Sentiment** words found in the text being analyzed that you may want to add. Here we added “expensive” to the dictionary assigning it a score of ‘-60’.

Term	Score	Possible Sentiment	Count	+Sentiment	Document Text
:(-70	salty	5	+90	[20] These chips are good but they're too expensive. The bag was half empty and it only took...
:)	70	many	4	+60	
:-(-70	expensive	3	+30	[88] ...chemicals and pesticides. The chips are a little more expensive than some others, but you can't beat the flavor...
:)	70	oily	3	-30	
:{	-70	overwhel...	2	-60	[128] So expensive! I know they're organic, but they're not worth...
:}	70	just	2	-90	
:d	80	enough	2		
:-d	80	delicious	1		
<3	60	crunchy	1		
7	70	sure	1		

Term	Score
exceptional	90
excite	60
excited	75
excites	60
exciting	75
expensive	-60
fabulous	90
fantastic	90
favorable	40

- You can also upload unique dictionaries by selecting **Manage Sentiment Terms** under the red triangle.
- You can choose to focus the sentiment analysis on select **Features** identified in the data.

Visit **Basic Analysis > Text Explorer** in **JMP Help** to learn more.